MMH,* Home Work 1 HW1

Elwannas HIRI†

December 31, 2020

Homework content In this homework we will try to answer some essentialy question to be able to write an article about *Build a parallel Corpora between Algerian Dialect and Modern Standard Arabic for Machine Translation*[1] and publish it.

1 Subject Area:

The field of the project is *Natural Language Processing.*

2 Aim:

Collect as much as possible for parrallel MSA-DA dataset of Algerian dialect.

3 Arguments:

There are many researchs about Arabic Dialect, they consist to LVA, EGY, MD and middle east dialects. in the other hand there are only a few reseachs of Algerian Dialect case. In addition to the diversity of Algerian Dialect. this is a good motivation to start working on this project.

---

*Methods Meta-heuristic

†Computer science student at Ahmed Draia University -Adrar-, github: https://github.com/t1g3r01

[1]MMH mini-project subject

4   Objectives:

- Filter collected data.

- Organize the dataset depends on location.

- Build a parallel dataset ready for the use for future works.